MINING AND MODELING DECISION WORKFLOWS FROM DSS USER ACTIVITY LOGS

Razvan Petrusel
Faculty of Economical Sciences and Business Management, Babeş-Bolyai University
T. Mihali Street 58-60, 400569, Cluj-Napoca, Romania

Keywords: Decision process mining, Decision workflow, Financial decisions, Process mining.

Abstract: This paper introduces the concept of decision workflows, regarded as the sequence of actions of the decision maker in decision making process. We show how, based on a decision support system we previously created, we log the behaviour of the decision maker. The log is then imported into ProM framework and mined using existent process mining algorithms. The mined model will show us the control-flow perspective (which is the order of decision maker's actions), the organisational perspective (which is the actual relationship among decision makers in group decisions), and the case perspective (what kind of support is required by each type of decisions). The aim of our research is to automate the creation of decision making patterns. Once obtained, the workflows can be merged into a financial enterprise model, which, properly validated, can become a financial reference model.

1 INTRODUCTION

There is a great variety of financial decisions among small and medium enterprises in Romania. We previously researched those decisions using a mixed approach based on questionnaires, direct observations and interviews (Petrusel, 2008). We evaluated financial decision making in over 50 small and medium enterprises in Romania. In the sample, we observed that decision makers use similar information sources and perform the same activities for similar decisions. This led us to believe that, up to a certain point, a decision process is actually very similar to a workflow. Our final target is to use decision workflows in order to create decision reference models. Those models will allow us to assess the quality of the decision procedures and the overall quality of decision making.

We rely on the work in process mining presented in (van der Aalst, 2002), (van der Aalst, 2004), (Wynn, 2008), etc. By using our approach, we argue that existing process mining algorithms can be employed in modelling decision making processes. We show how a DSS can be used to create a log regarding the decision makers' actions. This log will be mined using process mining algorithms in order to extract a decision making workflow. Several such workflows, extracted for similar decisional situation, can become a model of a decision making process.

We will present in this paper the first experiment on creating a financial decision model derived from workflows. We present in Section Two our research framework and several research questions. In the third section we will show how we used our DSS in order to create a log of the financial decision making activities. Then, we will show how we used process mining algorithms to create decision workflows. In Section Four, we will present our conclusions after this first experiment.

2 RESEARCH FRAMEWORK

There are several research questions that we try to answer in this section:

a) "Can enterprise financial decisions be treated as workflows?",

b) "Are there any tools for mining financial decision process models?"

c) "How can decision process models be used in order to create an enterprise model?",

Workflows are regarded as a depiction of the sequence of operations performed by an individual (Van der Aalst, 2002). A decision is an outcome of a cognitive process leading to the selection of an alternative from several possible choices. The

Razvan P. (2009).
MINING AND MODELING DECISION WORKFLOWS FROM DSS USER ACTIVITY LOGS.
In *Proceedings of the 11th International Conference on Enterprise Information Systems - Information Systems Analysis and Specification*, pages 144-149
DOI: 10.5220/0001987701440149

decision process implies sequential activities starting with the recognition of the need for a decision and ending with the choice of one alternative. The enterprise financial decisions involve several activities that the decision maker undertakes from the moment the need for a decision arises and until an alternative is chosen. We believe that these are enough reasons to argue that the decision process can be approached as a workflow.

The problem is that most of the decision related activities take place inside the mind of the decision maker. However, some of the mental activities need support and require additional information. But, in order to get that support, the decision maker interacts with the software, therefore leaving a trace regarding the sequence of his activities that can be logged. This is why our first goal is to create means to map decision processes to software tools, especially decision support systems.

The styles and methods of decision making may vary greatly among individuals because personal cognitive style will influence decision making (Martinsons, 2001). This is why we narrowed the research topic to decisions regarding the finances of the enterprise. We argue that, if decision making is seen as a workflow, for a specific set of decisions, comparisons between different enterprises and different individual decision makers are possible.

There are a lot of different software tools that help the creation of workflows and business process models. In order to create the process models, usually, there is a need for an expert that can examine the environment. There are also tools and algorithms that can be used to extract workflows from logs. But what if a log is not available because the decision making is done inside the mind of the user? And what if the use of experts requires a lot of time in order to get acquainted with the specific environment of each studied enterprise and this way the costs far exceed the benefits? We approach this problem by logging the actions of the decision maker while using a DSS. Based on the logged behaviour, we argue that a workflow can be created for each type of decision using existing process mining algorithms and software tools. This log is even more important in the case of a group decision.

The type of log that is mined determines what type of results will be available and which are the perspectives over the organisation that can be obtained (Dumas, 2005). By using process mining algorithms, we wish to gain some insights into the control flow of the activities perspective, and into the organisational perspective.

The control flow perspective gives insights into the tasks that are executed and the order of their execution (van der Aalst, 2002). It should also be possible to link the tasks in the model to the process instances performed.

The organisational perspective shows information regarding the social networks in an enterprise based on work transfer or on work subcontracting (van der Aalst, 2002). This is a very interesting perspective for our research because it gives a clear picture on who is the person with the most initiatives, who depends on other persons or who delegates the responsibility.

If the log has enough data, a case perspective can help improve the case-based forecasts regarding future decisions. And it can also enable the creation of a decisional profile for each decision maker.

The three perspectives discussed above relate to the most important questions in financial decision situations: “who?” (organisational perspective), “What?” (case perspective) and “How?” (control-flow perspective). Therefore, the general purpose statement of our research can be: “*Who* decides, on *what* decides and *how* is the decision made?”.

In order to answer those three questions we decided to use the ProM framework for process mining the logs obtained from CFAssist. The ProM framework is an open-source tool tailored to support the development of process mining plug-ins. This tool already contains a wide variety of plug-ins, some of them going beyond process mining (like doing process verification, converting between different modelling notations etc) (Verbeek, 2006).

Can current financial decision processes be reengineered by using new software and technologies? Is it possible to improve current processes? An approach to reengineering is to analyze current workflows and then try and improve them. How important is the possibility that for different enterprises the approach over decision making process is different? It is clear that, at international level, decisions have different premises (Martinsons, 2001). We argue that, an approach based on decision processes in the same region can be successful. We will start by analyzing decision making processes in our region. We will then create enterprise models for different views over the companies (the first one will be the financial view). We aim then to compare different workflows so that we have a better understanding of how decisions are made. If the occurrence of some decisional patterns is high, we can then propose them as reference models (or best practices) for certain types of decisions. For the companies interested in improving

their competitiveness this can be the best first step in reengineering their decision processes.

3 THE PROCESS MODEL

3.1 Creating the Log

The starting point of our research was the evaluation of enterprises in search for some kind of a log of overall operations. This log is supposed to be the foundation on which process models can be created. The conclusion is that Romanian small and medium enterprises do not use advanced ERP systems or any other software responsible for overall enterprise workflow management. The only exceptions are the companies that needed ISO certification. However, we found that even those companies do not use software to log all the activities for all the employees.

The first problem we faced was the need to create a log of actions regarding the financial operations, and, more specific, financial decision making. Our previous research was focused on creating a decision support system (CFAssist) based on cash-flows that aimed to improve financial decision making in small and medium enterprises (Petrusel, 2008). We improved this system so that every action of the decision makers was logged, giving us a raw source of data. We created a different version of CFAssist that presents all available decision tools (what-if analysis, scenarios, indicators, reports, expert systems) in different windows. If a user needs to use several tools he always must open new windows. We also created forms, menus, and buttons that aim to associate each mental process to an action. We logged the actions of the user while using CFAssist as well as the time stamps for each action. This gave us a fair idea regarding how much time the user spent viewing data, running what-if scenarios and simulations. The whole concept is based on the assumption that CFAssist is the only tool used in researching a decisional problem and choosing an alternative.

There were two challenges in improving the data sources:

a) the users worked with CFAssist so that their actions were limited by the software

b) the system extracted data from the accounting system of the enterprise, so not all the actions of the other actors involved in decision making were logged.

We created four tables that can be imported into ProM framework using ProM Import tool (it converts Access tables to MXML format of ProM logs).

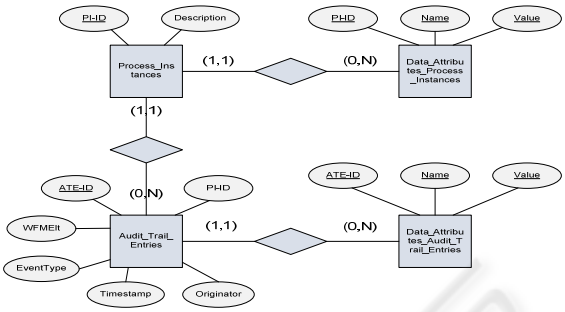

Figure 1: ER diagram for the four Process Mining tables.

In order to conduct our research we started with a test implementation in three enterprises that provided the training data. Because following the daily operations of the enterprises could take a long while and provide reduced relevance (since some strategic decisions are made rarely) we developed a list with nine detailed decisional situations based on each enterprise's data. For the first two enterprises there was only one decision maker while for the third enterprise there was the need for a group decision (two decision makers, each decision required consensus). The decision makers were required to make a decision based on each scenario. This provided us with three activity logs that could be used further in the mining process. The main scenarios were instantiated to suit the actual data known by the decision makers.

For example, one of the scenarios regarding the financing sources of investments was generally stated as follows: "the company decided to purchase a new car. The total value is <amount> euro and half of the total amount will be paid in advance and half will be paid on delivery (in two months). The decision alternatives are: finance from internal sources; bank credit; operational leasing; financial leasing; a mix of the previous sources; or drop the financing." Making a decision requires an evaluation of the financial position of the enterprise. For our study it is important which reports are used by the decision makers, which what-if analyses and scenarios are run, which indicators are selected for comparison and what is the final choice. The mined decision process model for the third company will be presented in 3.2 sub-section.

For each object in CFAssist we added code to insert data into the four tables as the decision makers used it. The actions of the user can be best seen in the tables Process_Instances, Audit_Trail_Entries and Data_Atributes_Audit_Trail_Entries. For the

scenario presented above, some of the records of the three tables are presented in Figures 2, 3 and 4:

PI-ID	Description
1	C1 S1 rank payment of accounts payable
2	C2 S1 rank payment of accounts payable
10	C1 S4 cashing method_bonuses
11	C2 S4 cashing method_bonuses
16	C1 S6 decide on financing sources to buy car
19	C1 S7decide on building acquisition
22	C1 S8 decide on car acquisition
25	C1 S9 decide on expansion

Figure 2: Some records in Process_Instances table.

ATE-ID	PI-ID	WFMElt	EventType	Timestamp	Originator
87	18	decision needed	start	24-Oct-08	D1
88	18	decision needed	start	24-Oct-08	D2
89	18	historic c-f	start	24-Oct-08	D1
90	18	c-f simulation	start	24-Oct-08	D1
91	18	historic c-f	start	24-Oct-08	D2
92	18	c-f simulation	start	24-Oct-08	D2
93	18	input data	start	24-Oct-08	D1
94	18	c-f simulation	start	24-Oct-08	D1
95	18	revenues and exper	start	24-Oct-08	D1

Figure 3: Some records in Audit_Trail_entries table.

ATE-ID	Name	Value
87	type test code	c3s6
88	type test code	c3s6
89	choose period: begins	january
89	choose period: ends	october
107	vote	financial leasing
108	vote	financial leasing
109	consensus	yes
112	decision: D1	execute
112	decision: D2	execute
112	decision: result	financial leasing
112	decision: send	e-mail

Figure 4: Some records in Data_Attributes_Audit_Trail _Entries table.

3.2 Importing the Log

After the log was obtained, a pre-processing was needed. This activity aimed to remove all data that is not necessary or that is incomplete. We consider that, a complete decision process starts with the detection of the need for a decision and ends with the choice of one alternative. Therefore, the start point of every process must be the "decision needed" task. This task is logged when the decision maker starts CFAssist and clicks the "Decision Support" button on the start-up form (as presented above). In the log tables along with the event is also stored the timestamp. The decision process is ended either with "communicate decision" or with "drop decision" tasks. Each task is logged when the user clicks either "send decision" button or "discard decision" button. All the actions of the user between those two tasks, logged as events for the objects in the systems, represent the decisional process. Incomplete processes that either do not start with "decision needed" task or are not ended with "communicate decision" or with "drop decision" tasks were removed from the log. In our simulated test environment there were only a couple of such processes. Even in real conditions we do not expect numerous such processes once the users get to know the system.

We used data filters when mining different decision processes. In order to mine the decision process of choosing the financing sources, along with the raw log we also used the filtered log by each of the decision makers. The filtered logs allowed us to create a separate decision workflow for each decision maker that can be compared.

3.3 Decision Process Models

After the logs were obtained and cleaned we used ProM framework in order to create the workflows. The main reason for our choice is that there are numerous plug-ins available that allow extensive mining and analyzes of the logs. Each plug-in gives us the opportunity to use a different algorithm to mine the available log.

We used alpha++, heuristic miner and fuzzy miner algorithms to order activities in the logs. The resulting models mined after using the three plug-ins were almost identical. This was caused by the fact that the test logs were almost noiseless due to the controlled test environment.

The order of decision making activities gives us a control-flow perspective over the decision process. The final goal is to establish dependencies among tasks. In decision making processes this means answering to several questions: which activity precedes which, are there any activities that imply others, are there concurrent activities (we observed that in decision processes concurrent activities usually means reviewing information from two sources) and if there are any loops (in decision processes we observed that loops appear mainly when what-if analyses and scenarios are reviewed). Another important piece of information is whether a path is more frequent than the others. If there is not a high frequency for one path it means that the user does not have a routine but searches for information in different places. This was found mainly in unstructured decisions that appear rarely (like strategic decisions and sometimes tactical ones). In operational decisions, the path is almost always the same. In this case, if the same path is followed by many enterprises, we can create some best-practice recommendations and have a base for a reference model. We will discuss two of the models obtained by using alpha++ algorithm on the logs filtered for the scenario presented in the previous sub-section:

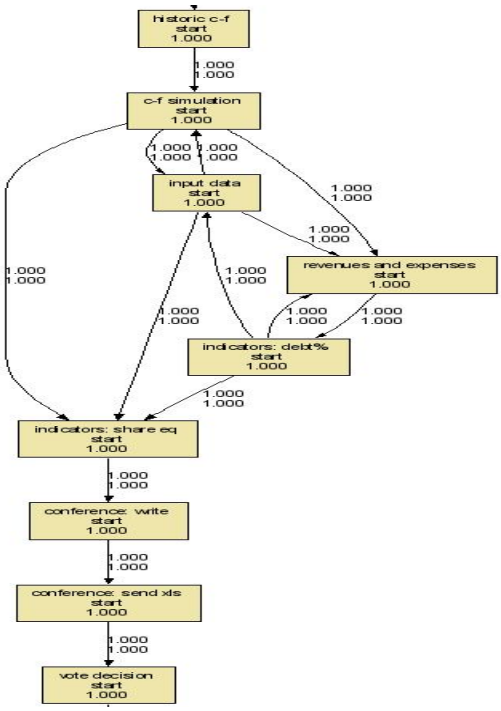

Figure 5: Partial decision workflow for D1.

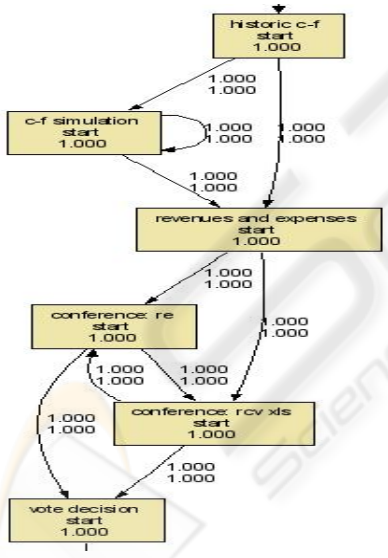

Figure 6: Partial decision workflow for D2.

Figure 5 and Figure 6 show a part of the decision workflow for the filtered logs. We discuss the decision analysis part because it better describes the strategies that are employed by the two users. It can be seen that D1 is more analytical and relies on more simulations and what-if analyses. By feeding new data into CFAssist he changes the initial values and tries to broaden his perspective over the decisional situation. D2 relies only on simulations based on accounting data and jumps to the decision without careful consideration. It also can be seen that D1 initiates the debate over the right decision and sends an Excel file to support his option. By following the two decision workflows we can argue that D1 has carefully considered all the alternatives and his choice is based on an analysis. Meanwhile, D2 briefly reviewed available data and jumped to the decision (possibly relying on experience). Even though the decision needs consensus, it can be argued that D1 influenced the final decision since he initiated a debate and sent a file to D2 in order to back up his choice.

In the decision workflows of the other scenarios the same trend could be observed. While there is no difference between the two decision makers from the point of view of former experience and studies (both have worked around eight years in similar positions and have graduated an economics faculty) it can be said that D1 is more involved in decision making and usually influences the other.

When we disclosed our findings regarding their decision profiles, both decision makers agreed that, in the majority of cases, the alternative suggested by D1 is the one chosen. Decision makers from the other enterprises also validated our decisional patterns as being close to reality.

With the extension of CFAssist, the creation of social networks became an important issue. In case of decision groups there is another important issue: “how are the communications between actors performed and what are the dependencies between the decision makers?”. This question can be answered by mining for social networks. In the test enterprises the first part of the question was relevant. But, in case the importance of each decision maker is not equal, or if the decision is taken in steps, at several management levels, the second question can also become important. Another thing that can be discovered in group decision making (especially where consensus is needed) is if there are any decision makers that rely on the opinions of other decision makers.

If the log exists and the process model was already created, the ProM framework allows the validation of the models by using conformance checker plug-in. After a process model is created it can be checked to see how much it matches existing execution data and to highlight discrepancies. The validation enables us to check how much a reference model is different from the actual decision process of an enterprise. Conformance checker can help us compare two decision models from different enterprises. We can determine the differences by

selecting one model as the prescribed model and checking it against execution data from the other enterprise. The points of non-compliance need to be examined in order to determine the differences.

An important point of interest is the Decision Point Analysis because it can lead to the discovery of Business Rules. By analyzing decision points we can determine the probability for a certain action to follow another action. This is an important factor when creating a reference model or when predicting the outcome of a decision.

4 CONCLUSIONS

This paper approached decision making process as a workflow. In a decisional workflow, all the actions of the decision makers are considered tasks that are sequential to one another. Our research interest covers the area of financial decision making in small and medium enterprises. The only way to trace the actions of the decision makers are by logging their actions while using software. One essential condition is to provide the decision makers with a tool that encourages the user to express all the personal decision making strategies. In order to do that, we modified a DSS we previously created so that all the actions of the decision makers while using the software are logged. Since the DSS was developed mainly around Microsoft Access we used ProM Import tool to convert Access tables to MXML logs. Those logs were opened in ProM framework. Using different available plug-ins we mined the logs for decisional process models and workflows. We analyzed the models and found that decisional models are comparable. We also argue that assertions can be made in connection to decision styles and strategies of different decision makers confronted with similar decisional situations.

The results obtained after our first tests are encouraging. We were able to compare financial decision making models obtained by mining logs from three enterprises. It is also relevant the fact that we could detect different decision making strategies and relationships in the case of group decisions. All decision makers involved in the experiment validated our profiles when we disclosed our findings.

Overall, we argue that we can approach decision making as a workflow and, that this approach can lead to decisional process models and patterns that can be compared. This comparison can improve perception of financial decision making in real-life enterprises and can be used as a base on which companies can reengineer their processes. The next phase of our research will aim to mine enough financial decision process models to create reference models for most common decisions in Romanian small and medium enterprises.

ACKNOWLEDGEMENTS

This research was founded through Grant type PN2 no. 91-049 / 2007 “Intelligent Systems for Business Decision Support (SIDE)”.

REFERENCES

Dumas, M., van der Aalst, W., M., P., ter Hofstede, A., H., 2005. Process-Aware Information Systems: Bridging People and Software Through Process Technology. John Wiley & Sons Inc New York.

Martinsons, M., G., 2001. Comparing the Decision Styles of American, Chinese and Japanese Business Leaders. In *Best Paper Proceedings of Academy of Management Meetings*.

O'Leary, D. E., Selfridge, P., 2000. Knowledge management for best practices. In *Communications of the ACM*, Volume 43, Issue 11, ACM New York.

Petrusel, R., 2008. A Decision Support System Taylored for Romanian Small and Medium Enterprises, In *ICEIS'08, 11th International Conference on Enterprise Information Systems*, INSTICC Press.

Van der Aalst, W., M., P., van Hee, K., 2002. *Workflow Management: Models, Methods, and Systems*, MIT Press Cambridge.

Van der Aalst W., M., P., Weijters, A., J., M., M., Maruster, L., 2004. Workfow Mining: Discovering Process Models from Event Logs. In *IEEE Transactions on Knowledge and Data Engineering, Volume 16 Issue 9*, IEEE Press New York.

Verbeek, H., M., W., van Dongen, B., F., Mendling, J., van der Aalst, W., M., P., 2006. Interoperability in the ProM Framework. In *Proceedings of the CAiSE'06 Workshops and Doctoral Consortium*, Presses Universitaires Namur.

Wynn, M., T., Dumas, M., Fidge, C., J., ter Hofstede A., H., M., van der Aalst W., M., P., 2008, Business Process Simulation for Operational Decision Support, In *LNCS 4928/2008*, Springer Heidelberg.